

Πρόταση Υλοποίησης για Πλατφόρμα Αυτοματισμού Δεδομένων

Προς: TechFlow Solutions

Από: Κωνσταντίνος Κωστόπουλος

Καλησπέρα,

Με βάση τις απαιτήσεις που συζητήσαμε σχετικά με την πλατφόρμα αυτοματισμού, παρουσιάζω την πρόταση μου, μαζί με μια εκτίμηση χρόνου και κόστους.

Προτεινόμενα Features και Χρονική Εκτίμηση

1. Εξαγωγή Δεδομένων από Φόρμες HTML

- Ο Ανάπτυξη parser για εξαγωγή στοιχείων πελατών από HTML φόρμες (2-3 ώρες).
- Ο Εξαγωγή πεδίων: Όνομα, Email, Τηλέφωνο, Εταιρεία, Υπηρεσία ενδιαφέροντος (1-2 ώρες).
- Ο Επικύρωση μορφής email και ελληνικών τηλεφώνων (1 ώρα).
- Ο **Χρόνος Υλοποίησης:** 4-6 ώρες

2. Επεξεργασία Email Files

- Ο Ανάπτυξη parser για EML αρχεία με αυτόματη κατηγοριοποίηση:
 - Εξαγωγή στοιχείων πελατών (1-2 ώρες).
 - Εξαγωγή στοιχείων τιμολογίων (1-2 ώρες).
- Ο Αυτόματη κατηγοριοποίηση emails (client inquiry vs invoice) (3 ώρες).
- Ο **Χρόνος Υλοποίησης:** 5-9 ώρες

3. Επεξεργασία PDF Τιμολογίων

- Ο Parser για HTML τιμολόγια με εξαγωγή:
 - Αριθμού τιμολογίου, ημερομηνίας (1-2 ώρες).
 - Ποσών, ΦΠΑ, συνόλου (1-2 ώρες).
- Ο Επικύρωση υπολογισμών (ΦΠΑ 24%) (1 ώρα).
- Ο **Χρόνος Υλοποίησης:** 3-5 ώρες

4. AI-Powered Extraction Engine

- Ο Ενσωμάτωση OpenAI GPT-4o-mini για έξυπνη εξαγωγή (1-2 ώρες).
- Ο Υπολογισμός confidence scores (2-stage approach) (3-4 ώρες):
 - Stage 1: AI confidence (70%) + completeness (30%)

- Stage 2: Validation penalties (-0.15 errors, -0.05 warnings)
- o Fallback σε rule-based parsing όταν confidence είναι πολύ μικρό (<70%) και το AI αποτυγχάνει (2-3 ώρες).
- o **Χρόνος Υλοποίησης:** 6–9 ώρες

5. Dashboard με Human-in-the-Loop

- o Ανάπτυξη Next.js dashboard με React + TypeScript (2 ώρες).
- o Σύστημα έγκρισης/απόρριψης/επεξεργασίας (3-4 ώρες).
- o Προβολή confidence scores με color-coding (πράσινο/κίτρινο/κόκκινο) (1 ώρα).
- o Φιλτράρισμα ανά τύπο (Forms/Emails/Invoices) (1 ώρα).
- o Analytics dashboard (Pending, Warnings, Approved, Rejected, Errors) (2-3 ώρες).
- o Προβολή source document (1-2 ώρες).
- o **Χρόνος Υλοποίησης:** 10–14 ώρες

6. Real-Time Updates με WebSocket & Redis

- o Ενσωμάτωση Redis για queue management και PubSub (3-4 ώρες).
- o WebSocket server για real-time ειδοποιήσεις (2-3 ώρες).
- o Frontend WebSocket client με αυτόματη επανασύνδεση (1-2 ώρες).
- o Toast notifications για νέα items (1 ώρα).
- o **Χρόνος Υλοποίησης:** 7–10 ώρες

7. Ενσωμάτωση Google Sheets

- o Αυτόματη εγγραφή εγκεκριμένων δεδομένων στο Google Sheets (4-5 ώρες).
- o Αρχικοποίηση και οργάνωση σε ξεχωριστά φύλλα (Clients/Invoices) και αρχικοποίηση headers κάθε στήλης (2-3 ώρες).
- o Retry logic για αποτυχίες API (1 ώρα).
- o **Χρόνος Υλοποίησης:** 7–9 ώρες

8. Logging & Monitoring

- o Structured JSON logging με structlog (1-2 ώρες).
- o Log rotation και error tracking (1-2 ώρες).
- o Health check endpoints (1 ώρα).
- o **Χρόνος Υλοποίησης:** 3–5 ώρες

9. Deliverables & Documentation

- o **USER_GUIDE.md:** Οδηγίες χρήσης για τελικούς χρήστες (2-3 ώρες)
 - Πλήρης περιγραφή λειτουργιών dashboard

- Παραδείγματα
 - Troubleshooting guide
 - Future features
- **SETUP.md:** Οδηγίες εγκατάστασης και configuration (1-2 ώρες)
 - Docker setup
 - Non-docker alternative setup
 - Google Sheets configuration
 - Environment variables
 - Troubleshooting
 - **README.md:** Τεχνική επισκόπηση έργου (1 ώρα)
 - Architecture overview
 - Technology stack
 - Quick start guide
 - **Διαγράμματα:** Architecture & Flow diagrams (3-4 ώρες)
 - System architecture diagram
 - Data flow diagram
 - Mermaid source files
 - **Τεχνική Πρόταση:** Αυτό το έγγραφο (3 ώρες)
 - Πρόταση υλοποίησης
 - ROI analysis
 - **Τεχνική Αναφορά (Report):** Ολοκληρωμένη τεχνική τεκμηρίωση (5 ώρες)
 - Πλήρης ανάλυση data flow με λεπτομέρειες
 - Architecture explanation με διαγράμματα
 - Cost savings analysis και ROI calculations
 - Technology stack justification
 - Testing strategy και coverage metrics
 - Scalability considerations
 - **Demo Video:** Επίδειξη λειτουργικότητας (5-10 λεπτά video) (4 ώρες)
 - Screen recording όλων των features
 - Script & Voiceover εξήγηση
 - Editing και post-production
 - Showcase πλήρους workflow (scan → review → approve → Google Sheets)

- Showcase error handling και logging προβληματικών αρχείων
 - ο Χρόνος Υλοποίησης: 19 – 22 ώρες
-

Σύνολο Χρόνου Υλοποίησης

64 - 88 ώρες.

Κόστος Υλοποίησης

- Ωριαία Χρέωση: 50€/ώρα.
 - Εκτιμώμενο Κόστος: 3,200–4,400€.
 - Λειτουργικά Κόστη (Μηνιαία):
 - ο OpenAI API (GPT-4o-mini): ~2€/μήνα
 - ο Google Sheets API: Δωρεάν
 - ο Redis: Δωρεάν (Docker local)
 - ο Σύνολο: ~2€/μήνα
-

Δοκιμές και Ποιοτικός Έλεγχος (Testing)

Η εφαρμογή περιλαμβάνει ολοκληρωμένο testing framework:

1. **Backend Testing (>80% code coverage)**
 - ο Unit Tests: Δοκιμή κάθε parser ξεχωριστά (Forms, Emails, Invoices)
 - ο Integration Tests:
 - End-to-end extraction flow με 25 dummy documents
 - Google Sheets integration testing
 - Redis queue operations
 - API endpoints testing
 - ο Property-Based Tests: Validation logic testing
2. **Frontend Testing**
 - ο Component Tests: Κύρια components (ExtractionCard, FilterBar, Dashboard)
 - ο Hook Tests: useWebSocket, usePendingQueue, useApproval
 - ο Integration Tests: User workflows (scan → review → approve)
3. **Automated Quality Checks**
 - ο Pre-commit hooks (Black, Ruff, pytest) για backend
 - ο Husky hooks (ESLint, Vitest) για frontend
 - ο Continuous testing κατά την ανάπτυξη

Χρόνος Υλοποίησης Testing: Ήδη ολοκληρωμένο (περιλαμβάνεται στις ώρες ανάπτυξης)

Επιπλέον Testing (Προαιρετικό):

- E2E testing με Playwright: 8-10 ώρες (320-400€)
 - Load testing: 4-6 ώρες (160-240€)
-

Προτεινόμενο Πλάνο Πληρωμών

1. 30% προκαταβολή για την έναρξη εργασιών.
 2. 40% μετά την ολοκλήρωση των βασικών λειτουργιών (extraction, dashboard, Google Sheets).
 3. 30% μετά την ολοκλήρωση και testing (final delivery).
-

Σχετικά με τις εργατοώρες, ο μέσος όρος 64 και 88 είναι 76, συνεπώς προτείνω να ορίσουμε τόσες ώρες στο πλαίσιο της συμφωνίας μας, στην οποία περίπτωση το τελικό κόστος εργατοωρών είναι **3,800€**.

Παραμένω στη διάθεσή σας για οποιαδήποτε διευκρίνιση.

Με εκτίμηση,

Κωνσταντίνος Κωστόπουλος

6974320060

costopoulos.constantinos@gmail.com